AMENDMENTS TO THE SPECIFICATION

Please replace the Title with the following Title rewritten in amendment format:

MULTILINGUAL ~~TEXT TO SPEECH~~ TEXT-TO-SPEECH SYSTEM WITH LIMITED RESOURCES

Please replace Paragraphs [0001], [0002], [0003], [0004], [0007], [0008], [0009], [0025], [0032], [0036] and [0040] of the Specification with the following Paragraphs [0001], [0002]. [0003], [0004], [0007], [0008], [0009], [0025], [0032], [0036] and [0040] rewritten in amendment format.

FIELD OF THE INVENTION

[0001] The present invention generally relates to ~~text-to-speech~~ text-to-speech systems and methods, and particularly relates to multilingual ~~text-to-speech~~ text-to-speech systems having limited resources.

BACKGROUND OF THE INVENTION

[0002] Today's ~~text-to-speech~~ text-to-speech synthesis technology is capable of resembling human speech. These systems are being targeted for use in embedded devices such as Personal Digital Assistants (PDAs), cell phones, home appliances, and many other devices. A problem that many of these systems encounter is limited memory space. Most of today's embedded systems face stringent constraints in terms of limited memory and processing speed provided by the devices in which they

are designed to operate. These constraints have typically limited the use of multilingual ~~text-to-speech~~ text-to-speech systems.

[0003] Each language supported by a ~~text-to-speech~~ text-to-speech system normally requires an engine to synthesize that language and a database containing the sounds for that particular language. These databases of sounds are typically the parts of ~~text-to-speech~~ text-to-speech systems that consume the most memory. Therefore, the number of languages that a ~~text-to-speech~~ text-to-speech system can support is closely related to the size and related memory requirements of these databases. Therefore, a need remains for a multilingual ~~text-to-speech~~ text-to-speech system and method that is capable of supporting multiple languages while minimizing the size and/or number of sound databases. The present invention fulfills this need.

SUMMARY OF THE INVENTION

[0004] In accordance with the present invention, a multilingual ~~text-to-speech~~ text-to-speech system includes a source datastore of source parameters providing information about a speaker of a primary language. A plurality of primary filter parameters provides information about sounds in the primary language. A plurality of secondary filter parameters provides information about sounds in a secondary language. One or more secondary filter parameters is normalized to the primary filter parameters and mapped to a primary source parameter.

[0007] Figure 1 is an entity relationship diagram illustrating a business model related to the multilingual ~~text-to-speech~~ text-to-speech system according to the present invention;

[0008] Figure 2 is a block diagram illustrating the multilingual ~~text-to-speech~~ text-to-speech system according to the present invention;

[0009] Figure 3 is a flow diagram illustrating the multilingual ~~text-to-speech~~ text-to-speech method according to the present invention;

[0025] The invention obtains the aforementioned results in part by using a system for an initial or primary language as a base. The quality of speech generated using this base in a second language is increased by a number of conversions from the secondary language to the primary language, and a number of extra units from the second language to be used in the synthesis. Given a speech unit as the basis for speech synthesis, the unit is separated into source and filter parameters and stored in memory. In general, the filter parameters provide information about the sound, and the source parameters provided information about the speaker. This source-filter approach is well known in the art of ~~text-to-speech~~ text-to-speech synthesis, but the present invention treats the two parts differently as can be seen in Figure 1.

[0032] Speech synthesizer engine 22 is adapted to convert text 24 from either the primary language or the secondary language to phonemes and allophones in the usual manner. The sound generation portion, however, uses both primary and secondary filter parameters with the source parameters to generate speech in the primary or secondary language. It is envisioned that a business model may be

implemented wherein a user of the device 14 may connect to a proprietary server 26 via communications network 28. Access control module 30 is adapted to allow the user to specify a selected secondary language 32, and receive secondary filter parameters 34 and a secondary synthesizer front end 36 over the communications network 28. It is envisioned that secondary filter parameters 34 may be preselected based on a priori knowledge of the primary language. It is also envisioned that the secondary synthesizer front end 36 may take the form of an Application Program Interface (API) that provides additional and alternative methods that may overwrite some of the methods of the speech synthesizer front end. The resulting multilingual ~~text-to-speech~~ text-to-speech system 38 may be adapted, however, to receive an initial set of secondary filter parameters and dynamically adjust the size of the set based on available memory resources of the embedded device.

[0036] Figure 2 illustrates some aspects of the multilingual ~~text-to-speech~~ text-to-speech system in more detail. Accordingly, system 38 has inputs 40 and 42 respectively receptive of text 24 and an initial set of secondary filter parameters 34. System 38 also exhibits speech synthesizer engine 22, source parameters 10, primary filter parameters 12, secondary filter parameters 16, mapping module 20, and normalization module 18 as described above. However, system 38 additionally has a similarity assessment module and memory management module 44. Module 44 is adapted to assess similarity of the initial set of parameters 34 to the primary filter parameters. Module 42 is further adapted to compare similarity of the initial set of secondary filter parameters 34 to a similarity threshold, to select a portion 48 of the secondary filter parameters 34 based on the comparison, to store the portion 48 of the

secondary filter parameters that are selected in a memory resource 46, and to discard an unselected portion of the initial set of secondary filter parameters 34. It is envisioned that the similarity threshold is selected to ensure that the secondary filter parameters 34 of the initial set that are related to sounds not present in the primary language are not discarded. It is also envisioned that module 44 may be adapted to monitor use of the memory resource 46 and to dynamically adjust the similarity threshold based on amount of available memory 50. Accordingly, system 38 is capable of generating speech 52 in multiple languages via an output 56 of the embedded device without consuming inordinate memory resources of the device in gaining the multilingual capability. The user of the device can therefore add languages as required.

[0040] Further, systems having important constraints regarding internal storage memory, can incorporate multiple language ~~text-to-speech~~ text-to-speech synthesis for the first time. In this case, a universal allophones to sound module is created with approximations to all possible sounds in all languages that need to be supported. The mapping from a particular language into the Universal set allows the generation of multiple languages with acceptable quality. Therefore, this invention provides an increase in value for products incorporating speech synthesis capabilities with a considerably small footprint in memory. This increase may have a great impact in mobile phones and PDAs, enabling the use of speech synthesis in multiple languages without memory constraints.

Please amend the Abstract section of the specification as rewritten in amendment format.